Bolt Beranek and Newman Inc.

(2)

LEVEL III

Report No. 4629

A093135
Dec 80

Command and Control Related Computer Technology: Packet Radio

Quarterly Progress Report No. 1
1 December 1979 to 29 February 1980

March 1981

Prepared for:
Defense Advanced Research Projects Agency

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER: 14 BBN [redacted]-4629	2. GOVT ACCESSION NO. AD-A096851	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): 6 COMMAND AND CONTROL RELATED COMPUTER TECHNOLOGY: PACKET RADIO.		5. TYPE OF REPORT & PERIOD COVERED: Quarterly Progress Report 1Dec79-29Feb80
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): 10 M. Beeler, J. Westcott, V. Strazisar		8. CONTRACT OR GRANT NUMBER(s): 15 MDA903-80-C-0206, new ARPA Order-2935
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Bolt Beranek and Newman Inc. 50 Moulton Street Cambridge, Massachusetts 02138		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 11 12 33
11. CONTROLLING OFFICE NAME AND ADDRESS: Advanced Research Projects Agency Information Processing Techniques Office 1440 Wilson Blvd., Arlington, VA 22209		12. REPORT DATE: Mar[redacted]81
		13. NUMBER OF PAGES: i., ii.,+ 28
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report): UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED. IT MAY BE RELEASED TO THE CLEARINGHOUSE DEPARTMENT OF COMMERCE FOR SALE TO THE GENERAL PUBLIC.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

9 Quarterly progress rept. no. 1 1 Dec 79-29 Feb 80

18. SUPPLEMENTARY NOTES

This research was supported by the Defense Advanced Research Projects Agency under ARPA Order No. 2935.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Packet Radio, Computer Communications, PDP-11, TCP, Gateway

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This document describes progress on development of a packet radio network. Activities reported include work on Station Software and Internetworking Research and Development.

DD FORM 1473 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

060100 [illegible]

BBN Report No. 4629 March 1981

COMMAND AND CONTROL RELATED COMPUTER TECHNOLOGY:

Packet Radio

Quarterly Progress Report No. 1

1 December 1979 to 29 February 1980

The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the Defense Advanced Research Projects Agency or the United States Government.

This research was supported by the Defense Advanced Research Projects Agency under ARPA Order No. 2935 Contract No. MDA903-80-C-0206

Distribution of this document is unlimited. It may be released to the Clearinghouse Deparment of Commerce for sale to the general public.

TABLE OF CONTENTS

1. INTRODUCTION

The Packet Radio station provides a variety of control, coordination and monitoring functions. BBN's role in developing this software is to specify, design, implement and deliver programs which perform these functions. BBN has additional roles in developing internet software for host machines and gateway software for interconnection of networks.

During this quarter several areas saw substantial progress. New commands were added to the Labeler, and conversion to PR network protocol (CAP) version 5.3, which supports 7-hop routes and affects several modules of station software, was made. These are covered in section 3, while documentation of the new Labeler is covered in section 2 along with other publications and negotiations.

Section 4 covers internetwork progress, particularly the successful operation of TCP Virtual Terminals and delivery of mini-gateway software for the LSI-11.

In section 5 we discuss progress in the design of transfer point routing and multistation operation, two areas in which we expect considerable implementation effort in the coming months.

Finally, section 6 presents hardware issues. Here the main points are installation of a Port Expander received from SRI, and installation of IPR units from Collins.

Also, during this quarter two problems arose which have received considerable attention. These are the TCP "data stream capture" problem (section 4) and the abundance of link quality changes, resulting in excessive Performance Data Packets sent by the PR units to the station to report these changes (section 3).

2. MEETINGS, TRIPS AND PUBLICATIONS

2.1 Meetings and Trips

BBN personnel attended the Packet Radio Working Group meeting at SRI December 4-6, where we participated in various discussions. Of particular interest was a presentation of the memory utilization in minigateways, as follows.

	PRNET minigateway	SATNET minigateway
MOS	3,700	5,400
BCPL	2,000	2,100
Gateway	10,900	13,000
	16,600	20,500
buffers	10,400	6,500
total	27,000	27,000

BBN personnel attended a PLRS/JTIDS/PR technical exchange meeting at Mitre January 16-17, and also the Internet meeting at SRI February 4-6.

2.2 Publications

PRTN 285, "Configuring the PDP-11/34 Station for Auto-Restart"

This note describes the minor modification which can be made to station PDP-11 hardware to cause it to automatically load and execute station software stored on disk, whenever power is turned on. This facility is desired at Fort Bragg, where power may temporarily fail while the station is unattended. Using this option implies loss of certain debugging capabilities, which this note explains.

CAP 5.3 "Labeler Operator's Guide"

This guide documents a set of eighteen commands available to the operator for modification and explanation of network status. The commands range from a printout of all link quality information received in PDPs, to a mini packet script of important control packets, to the ability to request information from PRs.

The commands are described under four major headings: General Command Format, PR Status, Network Connectivity Information, Labeler Control. These sections allow the operator to differentiate the effects of command types.

The guide concludes with an extensive example for new operators; all the commands are used.

We forwarded a draft version of this document to SRI and Rockwell International for comment and received several valuable suggestions which were included in the final document.

2.3 Negotiations and Informal Documents

2.3.1 Zeroing Link Qualities for Apparently Bad Hops

In the course of labeling a new PR, the Labeler only knows the connectivity to the new PR from the PR one hop away. The connectivity in the other direction is reported in a PDP from the new PR as soon as it is labeled.

Since the link quality information was incomplete, the Labeler zeroed the known one-way link quality if the Label packet failed. Our motivation was to reduce control traffic over inadequate links and we assumed that the bad link must be the one

on which we had no information. Otherwise, we would have received a link failure report from a PR. However, when a seven PR chain was put together to test the newly lengthened routes, the seventh PR was rarely labeled over the course of an entire day.

The problem stemmed from poor congestion control in the PRs which lead to so many packet collisions that LABEL packets, transmitted six times at each hop and three times from the station, (=18 tries over each hop) would frequently fail. Their failure resulted in the loss of link quality information from the PR which could no longer be reached, if any were present, and led to the problem of zeroed link qualities.

The zero link quality between the last hop and new PR is only corrected every half hour when a periodic PDP reports its view of network connectivity. The correction rate was too slow to allow labeling of the chain. Therefore, BBN removed the code which zeroed the link quality of the last hop over labeling paths which failed. This resulted in successful labeling of all seven PRs and in papers written on the congestion problem.

2.3.2 Transfer Points

We distributed a summary of the sub-session on transfer points and the TOP format from the December PRWG. Included was a warning that the design modification to PRTN 280 sacrifices certain functionality. This has sparked a series of negotiations between Collins, DARPA and BBN.

The two issues receiving the most attention are packet formats and whether transfer points should be able to forward traffic beyond the 4 hop routes.

A summary of the sub-session's resolution follows:

1. Our current design uses source/dest ID pairs rather than the source/dest PR pairs discussed in the paper. This will allow us to load split traffic. Currently, we do not do load splitting and only have 16 route table entries. Therefore, we are forced to have short optional timeouts in erasing old entries (30 seconds since last packet).

2. When the route fails we notify the station rather than the source.

3. We are not using route setup packets for the initial implementation. Rather, the station is sending out individual packets to each transfer point along the path (which is why #2 is necessary). It would ease the memory requirements in the station and reduce the volume of control traffic to allow the station to initiate its own route setup packet.

2.3.3 Multi-Station

Our multistation design has evolved considerably over time and become significantly different from the design documented in the November 1978 IEEE (find reference). BBN provided a brief comparison of the critical differences between the two approaches.

1. The previous design comparison did not keep track of station network connectivity. Our initial design will do so by a shortest distance matrix in which the distance metric is hop count.

2. We broadcast a request to all stations asking for the unknown destination's station to reply to the source station. Multiple replies will be stored in a cache and the route will be assigned to the nearest responding station.

3. The long, multiple station routes are assigned by each contributing station using the transfer point and forwarding mechanisms. This is radically different from the route finding packets which appear to wander through the subnet. It should provide significantly

better routes.

In general, the design has evolved to rely more upon the stations' perceptions of PRNET and station network connectivity. This places a heavier burden upon the stations but improves the quality of route assignment. Part of this is made possible by the advent of CAP5 which provides a greatly improved model of subnet connectivity.

2.3.4 Changes Needed for 7 Hop Routes

BBN agreed that it would be easier to use 7 hop routes at Fort Bragg than to implement a version of transfer points. We could then implement the full, stationless compatible transfer point design outlined in PRTN 256. That design solves our current forwarding problem by using source and destination PRs for routing. It also uses route setup packets for the establishment of long routes which is a significant improvement over the current scheme in which the station must contact each PR that holds a transfer point individually. We provided ARPA with the following summary of PRNET changes to support 7-hop routes.

Although everyone professes that the changes to their code are fairly small and self-contained, the sum total is not. We will be increasing the header length by three words and decreasing the text length by three words. The new route length enlarges the route table beyond the size able to fit into one packet. Therefore, the route interrogate command will have to be modified to notice whether the entire table has been returned and, if not, to ask for the remaining piece. There are also internet repercussions to the shorter text length.

Station Changes

XRAY, PRLOAD, CONN, and LABELER will all have minor changes to accommodate the proposed route length. In addition the LABELER must square its matrix another time - this must be done in any case. We do not have problems with the additional storage requirements.

EPR/IPR Changes

Collins has indicated that these will not be very difficult. The EPR would accept both header lengths to avoid changes in the PROM. As packets generated by the PROM are not forwarded by the receiving PR, a small test in the ROP handling code will be adequate. There are 100 free words; 48 will be taken by the longer route table. The first word in the command data packet returning routes will indicate whether the packet contains all the routes. A second command will be accepted to report the remaining routes.

TIU Changes

There will be small changes to the dispatcher, XRAY, the traffic source, PMON and TCP4. SRI estimates that two weeks would be required to complete all the modifications.

Subnet's Interface

The PR gateway will have a small change to accept the new header. The TENEX/TOPS-20 TCP4 will have to change a global variable decreasing the text length of TCP traffic by 3 words.

2.3.5 Exploring the Use of "C" in Place of BCPL

DARPA requested that BBN investigate the use of "C" in place of BCPL for station software. We have explored this possibility by analyzing the memory usage, code flexibility, and the choice of operating systems.

C is "known" to compile into better PDP-11 code but it does not have all the functionality of BCPL. Many improvements have been made to BBN's BCPL compiler so the differences in compiled code may not be so marked here as elsewhere. Since no known C-compiler produces PDP-11 code for an ELF operating system, we would have to operate under UNIX which does not support adequate process communications. Therefore switching languages would not be a simple transliteration.

The bulk of memory usage in the CONN process is in storage which would not be affected. CONN and Label take the most memory of the user processes. The number of SPP connections is controlled by storage requirements in CONN. New connections need additional packet buffers. If those were increased the next limitation would probably be interprocess ports - an ELF limitation. The major differences between C and BCPL are as follows:

1. Variable names are limited to 7 characters. Since much of our code uses mnemonics, 14 to 20 character names are common and their reduction to 7 would not be straight forward.

2. Everything must be declared in C; BCPL doesn't use declarations.

3. C employs a different pointer manipulation scheme. In part it uses declarations to decide how much to increment a pointer. It also forbids the movement of pointers within structures. Both of these restrictions impact existing code.

4. C doesn't have a VALOF-RESULTIS construct. We would use subroutines to achieve the same result.

2.3.6 Mixing SPP & SPP2

The Packet Radio community has raised the question of what would happen if we mixed two distinct transmission protocols. The motivation was to allow Collins to switch to a simpler protocol, SPP2, before work scheduled at BBN would permit the station to switch from SPP. We investigated the possibility that this mismatch might be transparent to station software and would allow the network to continue to function smoothly.

The mix of SPP and SPP2 is only transparent to the Station insofar as Station code remains the same. There is some cost of system inefficiency associated with mixing a full duplex and a simplex protocol.

The immediate effect will be seen in longer connection times within the Station. If the Station initiates a connection to the PR which does not require an answer, then the PR will not open its side of the SPP connection. In that case, the station must abort the connection. This occurs in improved labels, periodic labels, and some periodic route improvements. It should occur in all PTP assignments but the Labeler currently asks for a response which it then ignores.

The most obvious "critical path" problem would appear in a mobile PR which needs new routes and a new label. If the label is assigned and a CLOSE is sent, then a timeout must expire before the PTP routes may be assigned. This may add to congestion as alternate routing tries to handle the traffic and neighboring PRs report the problem in PDPs.

Whether the mixed protocol will work is dependent on the frequency of aborted connections and level of connection usage. The expected minimum is 3-6 aborted connections per 4 minute interval. The labeler has 8 specific SPP connections available. This may be enough.

We don't know the probability of these events or the seriousness of their impact on network performance. There may also be other, unforeseen problems.

The result of our research into these potential problem areas was a decision by DARPA to delay the Collins switchover until a new connection process could be written for the station.

2.3.7 Debugging Efforts

A serious congestion problem arises if 4 PTP routes, each generating traffic at 1/2 second intervals, are simultaneously erased from a PR. The resulting traffic forwarded through the station appears to block out control traffic and prevent routes from being assigned. A partial solution to this problem may be found by limiting the rate of traffic forwarded to the station by each PR. Traffic arriving every 5 seconds is adequate to permit PTP route assignment and will decrease the probability of congestion. This problem directly impacts the PR and CONN process.

Another problem SRI has experienced occurs during initial Labeler relabeling of an existing network. If there are a large number of PRs in the net and if there is a long chain of PRs which are not actively generating PDPs, the Labeler does not complete the network labeling in the first set of PDPs. Adjusting either the periodic PDP level or the speed in which unlabeled ("dead") PRs are thrown out of the tables should speed

labeling. Labeling was completed in the second round of PDP generation - 30 minutes from station startup. Delayed station labeling was corrected in this quarter by requesting a PDP from each PR immediately after Labeling.

During this quarter two issues arose, each concerning end-to-end (ETE) acknowledgement (ACK) packets. First, we suggested that software treatment of the transmit count field of these packets might be flawed, explaining the wildly varying link qualities seen at SRI. We negotiated with Collins and SRI on this, but a check of EPR, IPR and station implementation turned up no bugs.

Second, we heard at the PRWG meeting that PRs are supposed to accept certain ETE ACK packets with function field zero (non-SPP) as acknowledgements for SPP traffic the PR transmits. This seems completely wrong, so we asked for clarification. Since there was no response, we assume someone had misunderstood the protocol.

2.3.8 ARPANET-RCCNET Minigateway

We negotiated the use and installation plans for LSI-11 Port Expander serial number 21 with ARPA and other ARPA- sponsored interests at BBN. The tentative conclusion is that it will be used as an operational ARPANET-RCCNET minigateway, with a possible test period of a few days after installation. The ARPANET port will become available around May 1, so installation about then is planned.

3. THE PACKET RADIO STATION

3.1 Labeler

CAP5.2 Labeler debugging has been completed despite delays introduced by the testing facilities. The December PRWG meeting was very helpful in freeing an IMP port and providing access to SRI's facilities. Unfortunately, the conclusion of the meeting placed the station back on the port expander. This has not worked well. The major problem is that the station must be rebooted whenever the port expander breaks, which has effectively limited station access to SRI's work schedule and has required frequent communication seeking a rebooted port expander.

CAP5.3 testing has begun as a three step process. The first round, testing at BBN's two PR network, has been completed. This stage strikes out many of the initialization bugs, the packet format discrepancies, and the procedural misunderstandings between the PRs and the station. At the conclusion of this stage both station and PRs run for several days and the PRs are labeled.

The second stage of testing was at Collins rather than SRI. We have begun to use the Collins network because SRI's facilities are very heavily taxed. At Collins we can begin to test the point-to-point route assignment procedures in the station and the behavior of the PRs when routes are broken and created. This stage may reveal further discrepancies and provides excellent load testing for the PRs. However, testing at Collins, with three PRs and two TIUs, does not stress the station. Therefore the third stage of testing at SRI with a larger population of PRs is still critical to well-performing station software.

At SRI we have found a forwarding problem in the CAP5.3 design. We have proposed a fix to be implemented ONLY if packet forwarding must work to PRs beyond the 4 hop station perimeter. Once the forwarded packet reaches the transfer point PR (TFP), its source and destination fields will not allow the TFP to determine which route to put in the packet header. Hence, the TFP will request a point to point route from the station. We are awaiting comment as to the necessity of forwarding to distant PRs.

We are investigating SRI's report of bad labels appearing in the PRs. We have installed a switchable printout of all labeling activity which is coupled to a printout of PDP reasons. Since installation, 24 hours of operation under both light and heavy loads has revealed no bad labels - either in the PR or transmitted by the Labeler. However, the printout has revealed unexpected labeling activity which may be due to the Labeler's failure to communicate with a PR. SRI mailed the printout to BBN for appraisal which proved very helpful.

We have received two Labeler bug reports from SRI. One bug has been found and a solution is being coded. This was a race condition which had been in the Labeler since the beginning -- 3.5 years ago! It appeared now because of increasing stress the SRI testing is putting on the Labeler. The second bug is occasional bad labeling from the station. It had been seen to place FFFF preceding labeling assignments; the problem has been corrected.

Implementation of 7-hop routes in the station is complete and testing has begun. So far, we have established that XRAY, CONN and the Labeler can communicate with one EPR using 7-hop code. Testing of 7-hop PRLOAD (to EPRs) has been delayed until completion of some IPR work.

Slow communications pertaining to the remainder of CAP5.3 has led to delay as various contractors negotiate the contents. Further Labeler work will be required.

Modifications to Labeler Dialog Process

Three new Labeler commands have been added. "S"how Labels prints the text of label packets. "A"ssigned PTP Routes prints the text of new (or fixed) PTP routes. "U"pdated PTP Routes prints all routes improved during the periodic updating. Route printout includes source and destination IDs.

The Labeler printouts and command interface have been redesigned to increase readability and decrease paper usage by a factor of three.

The Labeler dialog process was modified to reduce the time period in which it hangs the routing process. This will result in PDP reasons, labeling assignments and other printouts occurring in the midst of normal typeout.

The Best Route command was modified to provide the previous best routes if the routing matrix is out-of-date. It will be indicated by an "(old)".

We implemented a new Labeler command which allows the operator to see SPP connection state changes. This has been valuable in tracking a connection handling bug.

The CAP 5.3 Labeler appears to have stabilized, although there are significant difficulties in labeling a 7-hop route due to gyrating link qualities. There are few moments when all link Q's on a 7-hop hardwired path are above 30 (out of 80 = perfect).

3.2 Performance Data Packet Issues

Spurious PDPs

The Station has been receiving many spurious link quality change PDPs. Some PRs appear to send them in once or twice a minute. This is a serious problem for three reasons:

1. It increases traffic in the network. Each PDP transmission results in six packets.

2. It increases processor load in the PR and, more seriously, in the Station. The station must open and close connections resulting in several signals between the labeler and CONN processes. In addition, the Labeler must process the PDP. While examining and storing the new link qualities is not very consuming of station resources, the reception of a PDP indicating that a serious link quality change has occurred invalidates the current routing matrix. Therefore, future route assignments cannot be made without recomputing the matrix.

3. It makes the recording of PDP reasons and (for current testing) the printing of labels transmitted very difficult. Frequently SRI must turn off the printout due to the enormous volume of paper it consumes. Most of this is used to print the reception of link quality change PDPs. It is certainly an odd problem to occur in network testing but that doesn't decrease its severity.

PDP Definition and Impact on Labeler Performance

It is true that the definition of a PDP could be expanded from the current one of reporting substantive connectivity changes. Theoretically speaking, a PDP could become the only control packet received from a PR. It could include TOPs and source/dest pairs for periodic updating, as well as suggested IPR fault messages.

However, this homogenization entails a substantial loss of information. If adopted, the Labeler would have to examine the PDP to see if the connectivity changes warrant updating the connectivity matrix. Looking at the PDP reasons is not sufficient as this PDP could carry a change of link Q from 7 to 4 without mention (a case normally covered by periodic updating). Therefore, we would find the Labeler pulling out old values for comparison - which is an inappropriate use of resources. To avoid unnecessary route computation (which is becoming more important as the length of routes and number of PRs increases), the handling of PDPs would have to be redesigned as the philosophy of their use is integrally bound into the code.

Rather than redesigning PDPs and their use, we propose that a new control data packet be created for the transmission of this information. Aside from the previously mentioned advantages, this would eliminate size restrictions. Network operators might find it valuable to receive an entire packet or internal table at the station. This information could be printed in specified formats and the Labeler could add comments for ease of operator interpretation.

3.3 7-hop Routes

Changes were made in several programs to support the new Packet Radio Net header format which was expanded to allow 7-hop routes. This change required modifications in the connection process, which implements SPP in the station, in the TCP resident in the station and in the mini-gateway. Changes to these programs were completed during this quarter and debugging in the Packet Radio Net at SRI was begun.

3.4 Support

3.4.1 Internet Bootstrap

The internet bootstrap used in PDP-11/35s and PDP11/40s to load these machines using the XNET debugger was modified to use 96-bit ARPANET leaders. This modification was necessary at this time in order to allow the Packet Radio Net station and mini-gateway to be attached to a port expander. Although the ARPANET IMPs still support 32-bit ARPANET leaders, the port expander supports only 96-bit leaders. All code run in the Packet Radio Net station and mini-gateway now uses 96-bit ARPANET leaders, so we are prepared for the eventual switchover to supporting only 96-bit leaders in the IMPs.

3.4.2 Port Expander

We hooked up the Port Expander supplied by SRI. We had misunderstood that the PE contained PROM memory; since it does not, a load line is needed to load software. Direct lines are needed because the loader program, "LSI", uses only direct line devices, not PTIP multiline controller (MLC) or TIP lines, the only other alternatives available. The only direct lines are on host BBND, which runs version 4 of the TOPS-20 monitor. We obtained a BBND direct line, but found that the loader program does not work under release 4. In fact, it crashes BBND. Some effort was made to determine the particular vulnerability which LSI exercises.

A test program was constructed which contains a sequence of monitor calls similar to that exercised by the LSI commands which crash the system. The test program also crashes BBND, but its sequence of monitor calls is still too complex to pinpoint the

problem. Testing is difficult because of the inconvenience to other users each time a crash occurs. Eventually facts arose pointing to a problem in certain site-specific monitor code installed by the BBN Research Computer Center staff. The problem has been turned over to them for resolution.

To continue PE testing, we borrowed a direct line on BBNA, which is not yet running the release 4 monitor. With this, we successfully loaded the PE. We were unaware, however, that the PE requires, for its console terminal, a 9600 baud, EIA terminal with no padding character requirements. Presently we are making do with a Tektronix 4023 terminal, which needs padding, and which is therefore difficult to use. We expect to arrange for a more appropriate terminal type in the near future. Because of these surprises about the support environment required by the PE, we have had to ask advice of SRI on several occasions; they have been cooperative and responsive, and PE installation is progressing.

3.4.3 Software Deliveries

Software was delivered to various sites this quarter as summarized in the following chart. "IMP11P" is the PDP-11/40 1822 interface (IMP11-A) diagnostic, configured to exercise the interface normally connected to the PR network.

software	Collins	SRI	BBN (cassette)
IMP11P	X		
CAP 5.2 station	X		
bootstrap	X		
EPR CAP 5.2.3	X		
EPR CAP 5.3.0	X		
EPR CAP 5.3.1	X	X	
IPR CAP 5.2.3			X
IPR CAP 5.3.0			X

4. INTERNETWORKING

4.1 Internet Protocol and Transmission Control Protocol

TOPS-20 release 4

The Transmission Control Protocol (TCP) version 4 and Internet Protocol (IP) version 4 have been integrated into both the straight DEC version and the BBN version of TOPS-20 release 4 sources this quarter. It assembles and loads, and is ready for testing. We are coordinating with the BBN Research Computer Center to schedule test time on BBNG. Also, the TOPS-20 internet user queue facility was modified this quarter to support single-port protocols such as the new XNET. This permits final debugging, and then use, of the IP4 XNET on BBNG.

TENEX

Progress was made this quarter on the TENEX version of the latest TCP. The system comes up, Network Virtual Terminal (NVT) and TELNET things work, and TCP Virtual Terminals (TVTs) work. Disturbances made to old code have been fixed except for one bug in the old user mode interface; TVTs do not use that interface, however, so they are not affected.

general debugging

We repaired an obscure bug in IP4 at ISI. A user tried to open and initiate a persistent connection to a foreign site which had no listening connection. This caused a logically correct, tight loop re-opening the connection each time a "reset" reply was received from the remote site. A delay was inserted to prevent the loop from usurping system resources.

During this quarter we also pursued the "data stream capture" bug reported by Fort Bragg users. This bug causes one user's TCP connection between Fort Bragg and ISI to receive good service, while other users' connections are apparently not serviced. We worked with SRI to get a handle on the problem. Then we constructed some files attempting to fix the problem and notified ISI of their availability. Later, we suspected that a particular subroutine call was clobbering the contents of an AC, and asked ISI whether this was so. The situation is complicated by some site-specific modifications made by ISI to the monitor and/or TCP they run. So far, the AC clobbering possibility is unresolved.

The TOPS-20 TCP version 4 is now relatively stable; we surveyed ISI, MIT and BBN, and found there are no known bugs which actually crash the system.

host gateways

We have begun coding to make host gateways take advice from internetwork gateways, in particular the output redirection capability.

4.2 Gateways

Experimentation with the SRI port expander pointed out a problem in the mini-gateway implementation. The mini-gateway declared an ARPANET host down, and would not send traffic to it for a period of two minutes, following the receipt of an incomplete transmission message from the network referring to that host. According to the 1822 interface manual, the incomplete transmission message is sent when there is a transient condition in the network which did not permit delivery of the

message. The port expander sends incomplete transmission messages to its attached hosts when it cannot deliver messages to the ARPANET because of some transient condition, such as lack of buffer space. As incomplete transmission messages are tied to transient failures in the network, it was incorrect for the gateway to assume that a host was down when an incomplete transmission message was received referring to that host. The mini-gateways were modified accordingly. The new version of the mini-gateway ignores incomplete transmissions and considers a host down, i.e., does not send traffic to it for some time, only after receiving a destination dead (type 7) message from the IMP or port expander.

5. TRANSFER POINTS AND MULTISTATION DESIGN AND NEGOTIATION

5.1 Transfer Points

BBN presented the transfer point design specificed in PRTN #280 ("Transfer Points") at the PRWG meeting in December. The design as specified was the minimal possible change from current routing that would support long routes and multistation operation. It had the shortcoming that forwarding by the station to a PR whose label was through a transfer point would not work. Consider this example:

```
A------S------B------C
```

In this example, C is labeled by station S through transfer point B. Assume A wishes to send a packet to C, but having no point-to-point route, sends it to S for forwarding. Since route entries in the PR are stored by source/destination device ID pair, S cannot forward the packet to C, as it is unlikely that B has a route entry for the source/destination pair A/C. Thus S would not be able to transmit or forward the packet until it set up a point-to-point route between A and C, or set up an entry for the pair A/C in transfer point B. At the time we wrote the PRTN, we decided that it would be better to have the inconvenience of this shortcoming than to specify a design (such as that specified in PRTN #256, "Stationless Compatible Routing") which would require a more substantial code change, since the design was only addressing a short-term need.

The PRWG meeting participants decided on modifications to the basic packet handling algorithm presented below.

When a PR receives a packet with an exhausted route it does

one of the following:

1. if the packet is destined for itself, process it
2. if the packet is destined for an attached device, send the packet over the interface
3. if there is an entry for the source/destination pair in the route table, write that route into the packet header and send the packet on its way
4. else, write the route to the station into the packet header and send the packet on its way.

The meeting participants decided to change step 4 to be "drop the packet and send a PDP to the station". This modification makes it impossible to forward any traffic through the station, even if no transfer points are allowed. Previously, when the station PR received such a packet the design specified that the packet be sent to the station. In the modified design, the station PR drops the packet and sends a PDP to the station. Not only is there the delay imposed by having packets dropped while awaiting a point-to-point route, but there is an additional delay imposed by the maximum rate at which PRs will generate PDPs, and the minimum interval in which the station will accept new connections from a given PR. Since this interval is approximately 20 seconds, point-to-point route assignment to a given PR is restricted to every 20-30 seconds.

BBN wrote and distributed minutes of the meeting, with an explanation of these problems. A flurry of network mail and phone conversations ensued. The first consensus was that transfer points should be implemented as originally specified in the transfer point PRTN. However, we also investigated how much trouble switching to 7 hop routes, versus implementing transfer points, would be. We were also asked to compare the PRTN's scheme with the full stationless compatible routing (SCR) scheme

specified in the SCR PRTN. The conclusion was that in the long term we really wanted the more *flexible routing* specified in the SCR PRTN, and that as a short term solution the 7 hop routes were preferable. This decision solved the immediate problems and was easier to implement.

The transfer point PRTN was written as the minimal possible scheme that would both allow longer routes and support multistation operation. The advantages of the scheme specified in the SCR PRTN over the minimal scheme are:

1. route renewal at any PR along the route

   This increases the power of alternate routing as there would always be several hops remaining in the packet header. However, this does require more processing by the PR, because all PRs along a route would have to search their route table entries, not just PRs at the end of route segments.

2. route tables per destination PR ID will replace source/destination device ID pairs

   This allows forwarding by the station to a PR labeled through a transfer point and reduces the number of route entries in the PRs. Looking at the change in stages, switching to destination only rather than source/destination pair will reduce the route count by a squaring factor. Switching to destination PR from destination will save by the number of devices normally assigned to a PR. Of course, since all possible routes are generally not assigned, we don't expect the real number of route storage slots to go down by large factors. And, unfortunately, source and destination PR IDs must be placed in the packet header.

3. route setup packets

   This makes the route setup procedure more efficient, as the station only needs to send a single packet to one endpoint of the route. The previous design required a separate packet from the station to every transfer point along the route.

5.2 Multistation Design

We designed a multistation capability, assuming transfer points as specified in PRTN 280, and documented the design in PRTN 281 ("Multistation Design Specification"). A meeting at BBN whose purpose was to discuss the design and to reach agreement among all contractors was scheduled for March 6, 1980. In preparation, we reviewed all the relevant PRTNs, studied the impact of implementing multistation based upon the routing scheme described in the SCR PRTN (replacing the transfer point PRTN), and organized the material for presentation at the meeting.

One impact of switching to the SCR PRTN's scheme is that packet header formats must be enlarged to include source and destination PR IDs (since route table entries would be indexed by destination PR ID rather than device IDs). Also, Route Setup Packets were not specified in the multistation PRTN, and their specification in the SCR PRTN did not include the multistation capability. Therefore we distributed a message specifying the modified packet header and Route Setup Packets.

Questions from Collins alerted us to the fact that our specification of error handling was confusing; it was specified differently in the multistation and the SCR PRTNs.

In the SCR PRTN, route failures are handled by issuing Route Loss Packets to the source PR. This would not always work, because packets from A to B might travel a different route than packets from B to A, and PRs along a route might not know how to get to the source PR. This was acknowledged in the PRTN, and other methods were also recommended to take care of the cases in which the source PR was unknown. However, with multistation, it would almost always be impossible to reach the source PR.

In the multistation PRTN route failures are handled by erasing the route as far back as the previous transfer point.

We recommended that the error handling scheme specified in the multistation PRTN be used.

We also noted that the Route Setup Packet would be more efficient if it set up the entire intersubnet route. This could be accomplished by sending the packet to the last PR along the route, and then to the station specified as "next station". We recommended that this modification be included in the multistation PRTN.

Finally, we have studied the impact of upgrading the transfer point design to support multistation and have specified packet formats which include all fields necessary to support both designs.

6. HARDWARE

At the conclusion of the PRWG meeting, we hand-carried and LSI-11 minigateway/Port Expander from SRI to BBN. The unit was transferred without connectors on the 1822 cables, since our use of a patch panel for the 1822 connections does not require the standard connectors. A slight confusion over signal assignments ensued, but was quickly resolved by help from SRI. Unfortunately, the DLV11-J four-line interface card was damaged when we accidentally connected a current loop console terminal instead of an EIA one. A temporary replacement has been borrowed and is working, permitting further checkout. SRI offered to repair the DLV11-J, and we are shipping it to them.

Collins personnel were at BBN December 12-19 to deliver and install IPR hardware, and deliver IPR software and documentation. We assisted in IPR checkout, and relatively successful operation was achieved. The EPRs were removed for relocation to Fort Bragg, and the unneeded two PR antennas and hardware which came with the IPRs were shipped back to Collins.

A few hardware problems occurred this quarter. The station disk drive broke and was repaired. The BBN PTIP port 076 broke during conversion to Distant Host mode (in support of minigateway/PE testing), and was repaired. And BBNA suffered some down time due to disk drive and monitor problems late in January.